3А

УДК 519.4

Д.С. Азаренко
Институт проблем искусственного интеллекта МОН Украины и НАН Украины, г. Донецк
Украина, 83050, г. Донецк, ул. Артема, 118б

Детектирование объекта на изображении и определение его смещения на двух различных изображениях

D.S. Azarenko
Institute of Artificial Intelligence, Ukraine
Ukraine, 83050, c. Donetsk, 118 "b" Artjoma st.

Detection of Object in the Image and Determine its Displacement in Two Different Images

Д.С. Азаренко
Інститут проблем штучного інтелекту МОН України і НАН України, м. Донецьк
Україна, 83050, м. Донецьк, вул. Артема, 118 б

Детектування об'єкта на зображенні та визначення його зміщення на двох різних зображеннях

В статье рассматривается задача поиска и детектирования объектов на изображении, а также определение разности положения объекта на двух различных изображениях. Для поиска используется детектор границ Кенни, а для обнаружения оператор Собеля. Так же рассмотрены методы определения положения объекта на изображении.
Ключевые слова: идентификация, детектирование, обнаружение границ, преобразование Хафа, метод оптических потоков.

The article deals with the problem of searching and detecting objects in the image and determine the difference position of the object on two different images. Search powered by Kenny edge detection, and detection Sobel operator. Also, the methods of determining the position of an object in the image.
Key words: identification, detection, edge detection, Hough transform, optical flow method.

У статті розглядається задача пошуку і детектування об'єктів на зображенні, а також визначення різниці положення об'єкта на двох різних зображеннях. Для пошуку використовується детектор меж Кенні, а для виявлення оператор Собеля. Так само розглянуті методи визначення положення об'єкта на зображенні.
Ключові слова: ідентифікація, детектування, виявлення меж, перетворення Хафа, метод оптичних потоків.

Введение

В статье речь пойдет об этапах поиска и детектирования объектов на изображении, а также определение разности положения объекта на двух различных изображениях.

Объект исследования: набор изображений движения объекта в пространстве. Объект состоит из соединенных сфер одинакового радиуса. Возможно смещение объекта по трем осям (x,y,z), а также произвольное вращение.

Предмет исследования: детектирование объекта на изображении, определение смещения (вращение) объекта на двух различных изображениях.

Цель работы – разработать алгоритмы, которые позволят детектировать объект на изображении и методы для определения изменения положения в пространстве объекта по двум различным изображениям. Совместно разработанные методы и алгоритмы будут представлять собой систему оптического распознавания движения объекта в пространстве.

Для достижения поставленной цели необходимо решить следующие **задачи**:

– поиск границ объекта;

– детектирование структурных элементов;

– определение положения объекта.

Поиск границ объекта

Края (границы) – это такие кривые на изображении, вдоль которых происходит резкое изменение яркости или других видов неоднородностей [1]. Проще говоря, край – это резкий переход/изменение яркости. Причины возникновения краёв:

– изменение освещенности;

– изменение цвета;

– изменение глубины сцены (ориентации поверхности).

Самым популярным методом выделения границ является детектор границ Кенни [2]. Основными этапами детектора границ Кенни является:

1. Сглаживание. Размытие изображения для удаления шума. Оператор Кенни использует фильтр, который может быть хорошо приближен к первой производной гауссианы $\boldsymbol{\sigma}$ = 1.4.
2. Поиск градиентов. Границы отмечаются там, где градиент изображения приобретает максимальное значение. Они могут иметь различное направление, поэтому алгоритм Кенни использует четыре фильтра для обнаружения горизонтальных, вертикальных и диагональных ребер в размытом изображении.
3. Подавление немаксимумов. Только локальные максимумы отмечаются как границы.
4. Двойная пороговая фильтрация. Потенциальные границы определяются порогами.
5. Трассировка области неоднозначности. Итоговые границы определяются путём подавления всех краёв, несвязанных с определенными (сильными) границами.

Детектор использует фильтр на основе первой производной от гауссианы. Так как он восприимчив к шумам, лучше не применять данный метод на необработанных изображениях. Сначала исходные изображения нужно свернуть с гауссовым фильтром.

Воспользовавшись оператором обнаружения границ (например, оператором Собеля), получается значение для первой производной в горизонтальном направлении (Gy) и вертикальном направлении (Gx).

Из этого градиента можно получить угол направления границы: Q=arctan(Gx/Gy).

Угол направления границы округляется до одной из четырех углов, представляющих вертикаль, горизонталь и две диагонали (например, 0, 45, 90 и 135 градусов). Затем идет проверка того, достигает ли величина градиента локального максимума в соответствующем направлении.

Например, для сетки 3×3:

– если угол направления градиента равен нулю, точка будет считаться границей, если её интенсивность больше чем у точки выше и ниже рассматриваемой точки;

– если угол направления градиента равен 90 градусам, точка будет считаться границей, если её интенсивность больше чем у точки слева и справа рассматриваемой точки;

– если угол направления градиента равен 135 градусам, точка будет считаться границей, если её интенсивность больше чем у точек, находящихся в верхнем левом и нижнем правом углу от рассматриваемой точки;

– если угол направления градиента равен 45 градусам, точка будет считаться границей, если её интенсивность больше чем у точек, находящихся в верхнем правом и нижнем левом углу от рассматриваемой точки.

Таким образом, получается двоичное изображение, содержащее границы (т.н. «тонкие края»).

На рис. 1 приведено начальное изображение объекта, а на рис. 2 – результаты детектора границ Кенни для заданной предметной области.

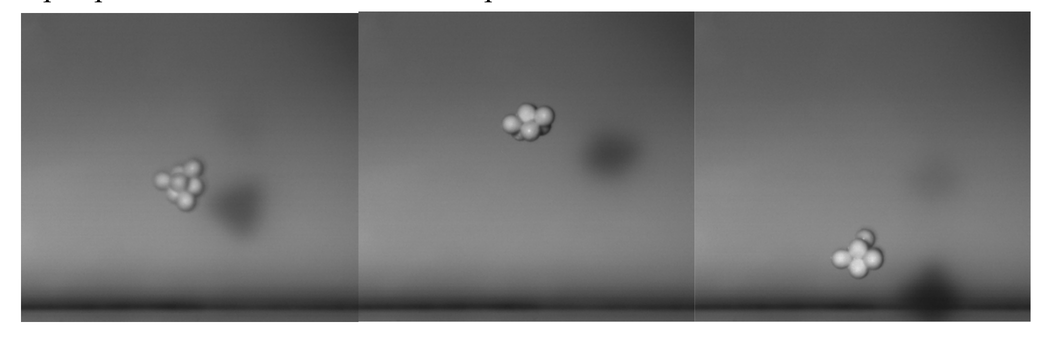


Рисунок 1 – Исходные изображения

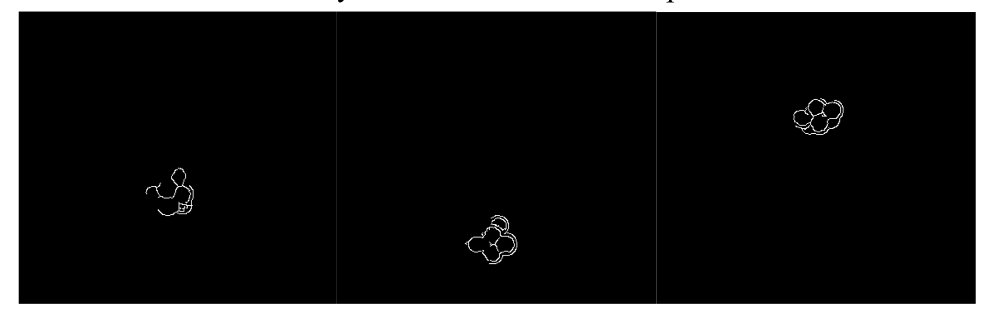


Рисунок 2 – Результат детектора границ Кенни

Детектирование структурных элементов

Так как объект на изображении состоит из окружностей, то по найденным границам необходимо восстановить окружности. Для структурного восстановления конструкции объекта. Один из наилучших для этого методов является преобразование Хафа [3]. Преобразование Хафа служит для поиска на изображении фигур, заданных аналитически: прямых, окружностей и любых других, для которых вы сможете придумать уравнение с небольшим количеством параметров.

В простейшем случае преобразование Хафа является линейным и использует уравнение прямой с угловым коэффициентом:

$$y = mx + b,$$

где m – тангенс угла наклона прямой к оси OX, b – координата точки пересечения прямой с осью OY.

Основная идея преобразования Хафа – учесть характеристики прямой не как геометрического места точек изображения, а в терминах ее параметров, т.е. m и b. Прямая представляется в виде точки с координатами (b, m) в пространстве параметров. Для каждой точки в пространстве (x, y) в пространстве параметров (b,m) будет соответствовать прямая, чем чаще через точку в пространстве (b, m) будет проходить прямая, тем больше будет коэффициент z у этой точки. Зная эти параметры, мы получаем положение линии.

Окружность описывается уравнением с тремя параметрами – $(x\text{-}x_0)^2 + (y\text{-}y_0)^2 = R^2$. Здесь (x_0, y_0) – координаты центра, а R – радиус. Пространство Хафа должно иметь три измерения, но мы ограничимся двумя, так как R нам известен. В этом случае все, что нам нужно найти – это координаты центров.

Возьмем такое исходное изображение (рис. 3). Так же, как и для преобразования Хафа для прямой, строим для каждой точки на исходном изображении точки в пространстве (x_0, y_0). Для точек с максимальным значением коэффициента z строим окружности на исходном изображении (рис. 4).

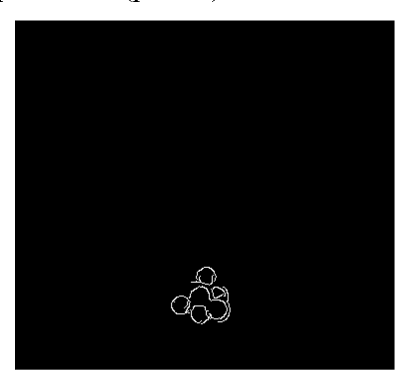


Рисунок 3 – Исходное изображение для преобразования Хафа

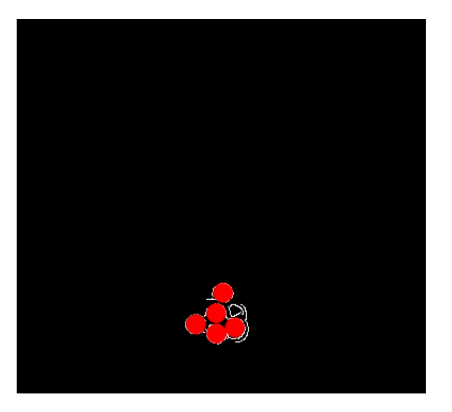


Рисунок 4 – Преобразование Хафа для окружностей

На полученном изображении видно взаимное расположение окружностей друг относительно друга, что нам и необходимо для следующего этапа.

Определение положения объекта

Для проведения этого этапа предлагается два подхода:

1. Распознавания объекта по построенному эталону;
2. Поиск изменений и смещений по кадрам.

Суть первого подхода заключается в том, что нам известна конструкция объекта, то есть взаимное расположение сфер относительно друг друга. Таким образом, мы можем смоделировать объект необходимой нам конструкции и производить над ним любые действия по перемещению вращению, сохраняя параметры изображения при этих действиях. Таким образом мы можем собрать базу данных изображений все возможных положений объекта в пространстве. При помощи этой базы производить распознавание [4]. При распознавании возможно использовать методы шаблонного сравнения [5], структурный подход или нейросети [6].

Шаблонное сравнение. Такой подход основан на сравнении исходного изображения со всеми шаблонами, имеющимися в базе; и выбирают шаблон с наименьшим количеством точек, отличных от входного изображения. Шаблонные системы довольно устойчивы к дефектам изображения и имеют высокую скорость обработки входных данных, но надежно распознают только те объекты, шаблоны которых им «известны».

Структурные системы. В таких системах объект описывается как граф, узлами которого являются элементы входного объекта, а дугами – пространственные отношения между ними. Системы, реализующие подобный подход, обычно работают с векторными изображениями. Структурными элементами являются составляющие объект окружности.

К недостаткам структурных систем следует отнести их высокую чувствительность к дефектам изображения, нарушающим составляющие элементы. Также векторизация может добавить дополнительные дефекты. Кроме того, для этих систем, в отличие от шаблонных, до сих пор не созданы эффективные автоматизированные процедуры обучения.

Нейросети. Рассмотрен еще один подход для распознавания на основе использования нейросетей. Преимущество этого подхода заключается в отсутствии необходимости предыдущего этапа с поиском границ и детектированием окружностей. Характерной особенностью неокогнитрона является возможность видоизменить сеть таким образом, чтобы она максимально соответствовала решаемой задаче. Опираясь на принцип функционирования сети, можно оптимизировать нейросеть для распознавания положения объекта, не меняя внутреннюю структуру, а только манипулируя слоями и количеством плоскостей в слоях. Их число зависит от количества эталонных образов и их характерных черт.

Неокогнитрон, рассматриваемый в данной работе, состоит из таких функциональных частей (рис. 5):

1) входной слой (рецептивный слой);
2) слой фильтрации входного образа;
3) слой усиления отфильтрованного сигнала;
4) промежуточные слои;
5) распознающий слой (слой классификации).

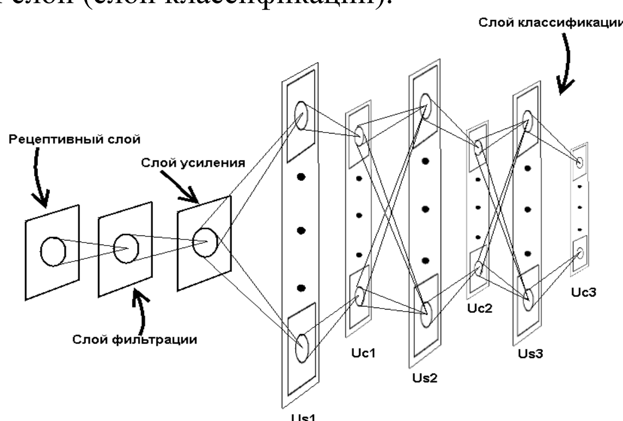


Рисунок 5 – Архитектура предложенной нейронной сети

Входной образ подается на рецептивный слой неокогнитрона, где он нормализуется и приводится к бинарному виду. За рецептивным слоем следуют слой фильтрации и слой усиления. Три слоя, рассмотренных выше, относятся к области первичной обработки сигнала, и их совместная работа позволяет «подготовить» произвольный образ к обработке.

Промежуточные слои состоят из попарно соединенных плоскостей простых и комплексных нейронов. В этой области проходит выделение основных свойств образов, а также сжатие информации для упрощения распознавания.

Последний слой – это слой классификации. От его структуры и настройки непосредственно зависит точность, с которой нейросеть может распознавать образы.

Недостаток подхода распознавания объекта по построенному эталону в целом является большой объем первоначальных данных. Необходимо построить объект, смоделировать его вращение, что займет много времени и ресурсов. А так необходимо обработать все полученные изображения при моделировании.

Второй подход представляет собой поиск изменений от кадра к кадру. На рис. 6 представлены два кадра верху оригинал внизу после поиска границ.

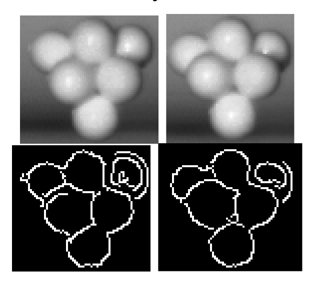


Рисунок 6 – Начальное и обработанное изображение

На каждом кадре после преобразования Хафа для окружностей будет найдено четыре четко выделенных окружности (рис. 7).

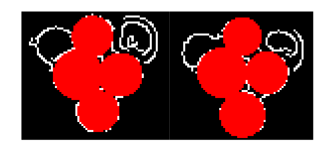


Рисунок 7 – Преобразование Хафа

Взаимное расположение окружностей за кадр не изменилось, но их положение в пространстве изменилось. По этим изменениям производится определение изменения положения объекта в пространстве. Своего рода данный подход основан на структурном распознавании, только структура не известна заранее, а строится для изображения и нас интересует изменение в структуре.

Схожим по функционалу является метод оптических потоков [7], который тоже может применяться для решения поставленной задачи. Оптический поток между парой изображений есть векторное поле, задающее естественную (в самом широком смысле) трансформацию первого изображения во второе.

Поле вектора оптического потока может быть представлено с помощью его компонентов $v_x(x, y)$ в направлении x и $v_y(x, y)$ в направлении y. Для измерения оптического потока необходимо найти соответствующие точки между одним временным кадром и следующим. При этом используется тот факт, что замкнутые участки изображения, сосредоточенные вокруг соответствующих точек, характеризуются аналогичными шаблонами интенсивности. Рассмотрим блок пикселов с центром в пикселе p, в точке (x_0, y_0), во время t_0. Этот блок пикселов необходимо сравнить с блоками пикселов, центрами которых являются различные потенциально применимые пикселы q_i с координатами $(x_0 + D_{x,}\, {}_y0 + D_y)$ во время $t_0 + D_t$. Одним из возможных критериев подобия является сумма квадратов разностей (*Sum of Squared Differences* – SSD):

$$SSD(D_x, D_y) = \sum_{(x,y} (I(x, y, t) - I(x + D_x, y + D_y, t + D_t))^2.$$

Здесь координаты (x, y) принимают свои значения среди пикселов в блоке с центром в точке (x_0, y_0). Найдем значения (D_x, D_y), которые минимизируют выражение для SSD. В таком случае оптический поток в точке (x_0, y_0) принимает значение $(v_x, v_y) = (D_x/D_t, D_y/D_t)$. Еще один вариант состоит в том, что можно максимизировать взаимную корреляцию следующим образом:

$$Correlation(D_x, D_y) = \sum_{(x,y)} (I(x,y,t)I(x+D_x, y+D_y, t+D_t))^2.$$

Метод с использованием взаимной корреляции действует лучше всего, если сцена характеризуется наличием текстуры, в результате чего блоки пикселов (называемые также окнами) содержат значительные вариации яркости среди входящих в них пикселов. Если же рассматривается ровная белая стена, то взаимная корреляция обычно остается почти одинаковой для различных потенциальных согласований и алгоритм сводится к операции выдвижения слепого предположения.

Допустим, что наблюдатель движется с линейной скоростью (или скоростью переноса) T и с угловой скоростью ω (таким образом, эти параметры описывают самодвижение). Можно вывести уравнение, связывающее скорости наблюдателя, оптический поток и положения объектов в сцене. Если предположить, что $f=l$, то из этого следуют уравнения

$$v_x(x,y) = \left[-\frac{T_x}{Z(x,y)} - \omega_y + \omega_z y\right] - x\left[-\frac{T_z}{Z(x,y)} - \omega_x y + \omega_y x\right],$$
$$v_y(x,y) = \left[-\frac{T_y}{Z(x,y)} - \omega_z x + \omega_x\right] - y\left[-\frac{T_z}{Z(x,y)} - \omega_x y + \omega_y x\right]$$

где $Z(x,y)$ задает координату z точки в сцене, соответствующей точке на изображении с координатами (x,y).

Достаточно хорошего понимания того, что при этом происходит, можно достичь, рассмотрев случай чистого переноса. В таком случае выражения для поля потока принимают следующий вид:

$$v_x(x,y) = \frac{-T_x + xT_z}{Z(x,y)}, v_y(x,y) = \frac{-T_y + yT_z}{Z(x,y)}.$$

Теперь становятся очевидными некоторые интересные свойства. Оба компонента оптического потока, $v_x(x, y)$ и $v_y(x, y)$, принимают нулевое значение в точке с координатами $x=T_x/T_z$, $y=T_y/T_z$. Эта точка называется фокусом расширения поля потока. Предположим, что мы изменим начало координат в плоскости $x - y$ для того, чтобы оно находилось в фокусе расширения; в таком случае выражение для оптического потока принимает особенно простую форму. Допустим, что (x', y') – это новые координаты, определяемые соотношениями $x'=x-T_x/T_z$, $y'=y-T_y/T_z$. В таком случае становятся справедливыми следующие уравнения:

$$v_x(x',y') = \frac{x'T_z}{Z(x',y')} v_y(x',y') = \frac{y'T_z}{Z(x',y')}$$

Литература

1. Сирота А.А. Статические алгоритмы обнаружения границ объектов на изображении / А.А. Сирота, А.И. Соломатин // Вестник ВГУ. – 2008. – № 1.
2. Буй Т.Т.Ч. Анализ методов выделения краев на цифровых изображениях / Т.Т.Ч. Буй, В.Г. Спицын [Электронный ресурс]. – Режим доступа : http://www.tusur.ru/filearchive/reports-magazine/2010-2-2/221.pdf

3. Дегтярева А. Преобразование Хафа / А. Дегтярева, В. Вежневец // Компьютерная графика и мультимедиа. – 2003. – Выпуск № 1(2).
4. Сойфер В.А. Компьютерная обработка изображений. Часть 2. Методы и алгоритмы / Сойфер В.А. // Соросовский образовательный журнал. – 1996. – № 3.
5. Метод быстрой корреляции с использованием тернарных шаблонов при распознавании объектов на изображениях / Глумов Н.И., Мясников Е.В., Копенков В.Н., Чичева М.А. // Компьютерная оптика. – 2008. – Т. 32, № 3. – С. 277-282.
6. Каллан Р. Основные концепции нейронных сетей / Каллан Р. – Москва : Издательский дом «Вильямс», 2001. – С. 288.
7. Круглов В.В. Искусственные нейронные сети : теория и практика / В.В. Круглов, В.В. Борисов – Москва : Горячая линия-Телеком, 2001. – С. 382.
8. Конушин А. Слежение за точечными особенностями сцены / А. Конушин // Компьютерная графика и мультимедиа. –2003. – Выпуск № 1(5).
9. Расторгуев А.С. Реализация субпиксельного уточнения ViFlow метода поиска оптического потока / Расторгуев А.С. [Электронный ресурс]. – Режим доступа : http://se.math.spbu.ru/SE/YearlyProjects/2010/YearlyProjects/2010/445/Rastorguyev_report.pdf

Literatura

1. Sirota A.A. Solomatins A.I. Algorithms to detect static edges of objects in the image. - Herald of the Voronezh State University, 2008, No 1
2. Bui T.T.CH., Spitsyn VG Analysis of methods for edge detection in digital images. http://www.tusur.ru/filearchive/reports-magazine/2010-2-2/221.pdf
3. Degtyarev A., Vezhnevets B. Hough transform. Computer graphics and multimedia. Issue number 1 (2), 2003.
4. V.A. Soifer. Computer image processing. Part 2. Methods and algorithms. - Soros Educational Journal number 3, 1996.
5. Glumov N.I., E.V. Myasnikov, V.N. Kopenkov, M.A. Chicheva. Rapid correlation method using ternary template for recognition of objects in images. Computer Optics, Volume 32, № 3, 2008, pp. 277-282
6. Callan R. Basic concepts of neural networks. - Moscow: Publishing House 'Williams', 2001. S. 288.
7. Kruglov V.V., Borisov V.V. Artificial Neural Networks: Theory and Practice. - Moscow: Hotline - Telecom, 2001. - S. 382.
8. Konushin A. Tracking the point singularities of the scene. Computer graphics and multimedia. Issue number 1 (5). 2003.
9. Rastorgouev AS The implementation of sub-pixel refinement ViFlow search method of optical flow. http://se.math.spbu.ru/SE/YearlyProjects/2010/YearlyProjects/2010/445/Rastorguyev_report.pdf

RESUME

D.S. Azarenko

Detection of Object in the Image and Determine its Displacement in Two Different Images

This paper discusses an approach to image processing, detection and edge detection of objects, detection of objects using the Hough transform and neural networks, as well as the definition of changing the position of an object in space by means of optical flow method. For the detection and edge detection in the course of the experiment was chosen edge detection operator Kenny and Sable, as the most suitable for the task. For the detection of structural elements of a Hough transform is used as an object composed of spherical elements. Determine the object in the image considered several approaches. Pattern search advantage is the ease of implementation. Template system is quite resistant to image defects and have a high processing speed of the input data but reliably detect only objects, templates which they "know."

A structural approach to the shortcomings which include the complexity of machine learning and resistance to image defects.

The advantage of using neural networks is the lack of need for image pre-processing, such as search boundaries, filtering and detection of structural elements.

The most promising solutions to this problem is the method of optical flow, which is devoid of a pattern and lack of structure search, and not as complex as a neural network. The quality of this method depends on the quality of the detection and the detection of the structural elements of the object. But for this method this step is not necessary.

Статья поступила в редакцию 05.07.2013.